Interview with Stefan Wrobel on "Applied Big Data Research"

Stefan Wrobel is Professor of Practical Computer Science at Bonn University and Executive Director of the Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS in Sankt Augustin, Germany. After computer science studies in Bonn and at Georgia Tech and his doctoral degree from the University of Dortmund, he worked as co-founder and Technical Director of the software house Dialogis in Bonn and as Professor of Knowledge Discovery and Machine Learning at the University of Magdeburg, prior to joining Fraunhofer. As an internationally known researcher in machine learning, he was among other tasks General Co-Chair of the ICML conference in 2010. Since 2013, Stefan Wrobel is coordinator of the Fraunhofer Big Data Alliance, an interdisciplinary association of 25 applied research institutes founded under his leadership.

Interview by

Prof. Dr. Matthias Jarke (✉)
Lehrstuhl für Informationssysteme
RWTH Aachen University
Ahornstr. 55
52074 Aachen
Germany
jarke@dbis.rwth-aachen.de

Published online: 2014-07-11

Prof. Dr. Stefan Wrobel
Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS
Schloss Birlinghoven
53754 Sankt Augustin
Germany
stefan.wrobel@iais.fraunhofer.de

BISE: Professor Wrobel, at the international information technology fair CeBIT in March this year, the Fraunhofer Society officially launched its Big Data alliance which now comprises 25 institutes. Just why is Big Data such an important topic for Fraunhofer?

Wrobel: At Fraunhofer, our research is driven by the needs of companies in the marketplace, and here we have seen a very big change across the past few years. While the fundamental trends that lead to the availability of more and more data have been developing for a longer period of time, we are now at a point where companies are very intensively realizing that the proper use of data can be a decisive factor for their competitiveness and can even help create revenue in entirely new business areas. But exactly how to put this into practice – given complex questions ranging from business strategy to technical implementation – seems to be very difficult for many companies, especially if they are smaller, so we are seeing an intensive demand for Fraunhofer here.

BISE: So in your view, what once was a hype topic seems to have finally arrived in companies for good?

Wrobel: From what we see, the potential of Big Data is recognized by a large majority of companies, and many recent studies seem to be confirming that this is so. In a study we performed ourselves for the German Ministry of Economy, 69 % of companies said they were trying to gain strategic advantages from Big Data, and other studies by MIT, BITKOM and others are arriving at similar percentages. However, it is also quite clear that the path from seeing the potential to actually realizing it is not an easy one. In our study, only 8 % of companies said they did not see major obstacles, and the other studies also show consistently that only relatively few companies have successful strategic Big Data initiatives in place.

BISE: Where do you see the key challenges that companies are facing when trying to reap the opportunities of Big Data?

Wrobel: A central point is certainly that Big Data is not only an IT topic. At first sight, Big Data is about transporting, storing and analyzing large volumes of data, all of which appears to be addressed by distributed storage and in memory analytics technologies readily offered in the marketplace. In practice, however, it soon turns out that Big Data projects must be oriented end-to-end towards the business value that one is seeking to generate. This often requires the cooperation of multiple stakeholders and a strategic new look at a company's data assets in combination with in-depth knowledge of the respective business sector.

BISE: Are you saying then that technical issues play only a minor role if any in Big Data projects?

Wrobel: Quite to the contrary – the point is, however, that they should be considered after the strategic business goals have been identified. Then it is certainly no trivial task to pick the right solutions from the large array of software available in the marketplace, and to combine and configure them in a tailor-made way for the individual needs of the company and its goals. I do see significant need for further research and innovation in fundamental Big Data technologies, in data management and in analytics. We need improved formal business models for Big Data processes, and we need to considerably improve the handling of the linking and analysis of heterogeneous data sources. The increased availability of streaming data especially in an industrial context pushes the boundaries of current algorithms, and the multimedia revolution with images and video is going to be the next fundamentally new Big Data challenge. And there is enormous potential in further improved privacy-preserving analytics because companies want to comply with the legal and ethical expectations of society without compromising their business goals.

BISE: Why do you believe that the Fraunhofer alliance Big Data will help companies address these challenges?

Wrobel: As stated above Big Data projects often cut across company functions and do not only involve technical, but also business and application sector know-how. No single Fraunhofer institute spans all these aspects. So we formed the alliance Big Data to give companies a central point of contact and make it easier for them to work with just the right combination of Fraunhofer competences. Our solution portfolio and technology stack range from data management to advanced analytics, but our alliance includes institutes focusing on the business aspects and, perhaps most importantly, also institutes with know-how in all the important application sectors – ranging from business and finance, energy and environment, life sciences and healthcare, logistics and mobility, production and industry 4.0, to security.

Some project examples: Big Data algorithms are monitoring billions of credit card transactions against fraud. Experimental data are analyzed to find cures for cancer. Sensor data are being processed for smart grid prognostics. Logistic chains are managed using large volumes of RFID data. Production processes minimize downtime and energy consumption, and crowd sourcing messages are analyzed to warn quickly in the event of an emergency.

BISE: How can we address the apparent lack of specialists with Big Data expertise in the short and long term?

The institutes are offering companies a program of courses offsite as well as onsite that address topics all the way from introductory Big Data issues to specialized technologies. Companies make use of this to train people within a compact timeframe who are going to be involved in Big Data projects. In addition, many of us are of course involved in setting up data science classes and curricula at our respective universities. And finally, we also cooperate with national and international partners to move Big Data and Data Science ahead in Germany and in Europe, for example with the Smart Data Innovation Lab in Karlsruhe, as well as with initiatives at the European level, e.g. by means of the technology platform NESSI.

BISE: Professor Wrobel, thank you for the interview!